

The axiom system of classical harmony

András Tóbiás

Currently at Technische Universität Berlin, Berlin, Germany
{tobias_AT_math.tu-berlin.de}.

This work was made at Budapest University of Technology, Budapest, Hungary.

Abstract. This paper constructs a new mathematical axiom system for classical harmony, which is a prescriptive rule system for composing music, introduced in the second half of the 18th century. The clearest model of classical harmony is given by the homophonic four-part pieces of music. The form of these pieces is based on the earlier four-part chorale adaptations of J. S. Bach. This paper gives a logical structuring of the musical phenomena belonging to the research area of classical harmony. Using these, it constructs a consistent mathematical axiom system, which incorporates the usual classical compositional principles as constraints for compliance with classical harmony, giving restrictions on the general solution set of homophonic four-part pieces.

Keywords: equal-tempered piano, trichotomy of keys, fundamental theorem of tonality, compositional principles, chord-changing constraints.

The main goal of this article is to provide a mathematical axiomatization for the strictly homophonic four-part model of classical harmony. This model is mainly based on J. S. Bach's four-part chorales, collected in [1], which had been written before the start of Viennese classicism. Later, in the second half of the 18th century, almost unequivocal, prescriptive compositional principles were determined for this four-part model. By their nature, these compositional principles form a mathematical axiom system, as soon as all the basic notions of music corresponding to classical harmony are mathematically well-defined. In this paper, we present a possible way of logical ordering of these musical notions, and having these definitions we present a *consistent and complete* axiom system which tells how to write homophonic four-part pieces.

Homophonic four-part pieces are interpreted as special right-continuous functions $M : \mathbb{R}^+ \rightarrow K^4$, where K is an *equal-tempered piano*. Here the real half-line refers to time. The special properties of M are that for any $t \in \text{Dom } M$, $M(t)$ is a special kind of *chord* in a special four-part version, and that the four voices always change their tones in the same time, yielding a *chord change*. The chords in the ranges of the homophonic four-part pieces have to be preliminarily given kinds of *triads* and *seventh chords* associated with a musical *key*.

The formulation of the axiom system is strongly related with constraint programming. The axioms of classical harmony, which are called *compositional principles*, determine about any homophonic four-part piece whether it complies with

classical harmony or not. One of the first axioms, the so-called *correctability condition* describes when a homophonic four-part piece complies with classical harmony *apart from the chord-changing points*. Our *fundamental theorem of tonality* (Theorem 1) gives a useful equivalent condition for correctability of feasible pieces on a finite time interval. It claims that among these pieces, the correctable ones are exactly the ones which are *locally tonal*. Thus, the usual chord-changing rules of classical harmony can be embedded in the axiom system as constraints. A correctable piece complies with classical harmony in a given chord-changing point if none of these constraints is violated. Hence, until it does not cause inconsistency, one can add or remove chord-changing rules to or from the axiom system according to new results of music theory. This virtually makes exercises of writing a correct four-part piece consisting of given chords a constraint programming problem. Some chord-changing compositional principles are mathematically described in e.g. [2, Section 5.11], but without aiming to constructing an axiom system with all connected notions mathematically defined.

The axiomatization is completed by the *modulation* (musical key change) rules, which give constraints on several consecutive chords which a modulation consists of. The modulation rules are musically quite complex but mathematically less interesting; an overview of them can be found in Section 6.

As for the compositional principles not detailed in [2], we follow the traditional Hungarian music theory coursebook [3]. In this article we use the German notation of classical harmony, according to the Hungarian convention, but we note that in the meaning of the axioms this makes no difference from e.g. the British notation system.

The contribution of this paper is to give a logical basis for musicians for writing a new music theory coursebook for high schools. This seems to be necessary in Hungary, and it can also be helpful in other European countries. In particular, we give a classification of musical keys. We show that there are *exactly three different keys* with the common base tone on the equal-tempered piano, up to enharmonic equivalence. The main theorem of this paper is the fundamental theorem of tonality, which makes it possible to embed the chord-changing compositional principles as constraints to a mathematical axiom system.

The paper is organized as follows. Section 1 enumerates and logically orders basic musical notions. It also explains compliance of triads and seventh chords with classical harmony. In Section 2, we define and classify musical keys. In Section 3, we present the model of homophonic four-part pieces. Musical functions and functional tonality are defined in Section 4, while the fundamental theorem of tonality and the structure of our axiom system is presented in Section 5. Finally, Section 6 summarizes the compositional principles for modulations.

1 Basic notions of music theory used in classical harmony

We provide an axiom system for composing homophonic four-part pieces of music, in a first-order language. We do not construct a new language but use the one of set theory, assuming the ZFC axiom system. We use simple physical prop-

erties of the overtone system, but formally these only have arithmetic meaning. As usual in music theory, the *tone* Y is a longitudinal wave moving in an elastic medium with frequency $f(Y) > 0$. For a tone X with frequency $f(X) > 0$, X is *audible* if $20 \text{ Hz} < f(X) < 20\,000 \text{ Hz}$. When speaking about tones, we always mean that the tone consists of all of its *overtones*. The set of overtones of the tone X is $\{Y \mid Y \text{ is a tone, } \exists n \in \mathbb{N}^+ : f(Y) = n f(X)\}$. The overtone of X with frequency $n f(X)$ is called the n th overtone of X . Thus, when we consider *a tone* X *with frequency* $f(X)$, it can be handled e.g. as $X = (f(X), f(X)) \in (\mathbb{R}^+)^2$, thus the definitions of this paper can be derived consistently from ZFC.

In the whole article, $B_r(x)$ denotes the open ball with radius r around the point x in any metric space, further \bar{A} the closure of A and ∂A the boundary of A in any topological space. Int denotes interior, Dom domain and Ran range.

Musical *intervals* are equal distances in the (base 2) logarithmic frequency scale. The most important intervals can be derived from the overtone system, cf. [2, Section 4.1]. The interval of a tone and its 2nd overtone is called *perfect octave*, the one of the 2nd and 3rd overtone of a tone *perfect fifth*, the one the 3rd and 4th overtone of a tone *perfect fourth*. We say that a tone X is higher than a tone Y (and Y is lower than X) if $f(X) > f(Y)$. Intervals can be summed, hence one can speak about *octave-equivalent* tones X and Y , the interval of which is n octaves with $n \in \mathbb{Z}$. If X is n octaves higher than Y , this means $f(X) = 2^n f(Y)$. Consequently, octave equivalence is an equivalence relation on the set of tones. The *octave-equivalence class* of the tone X will be denoted by $[X]$.

The following definitions will be used in Section 4, where we define musical functions and tonality. Let X be a tone and Y its 3rd overtone. The *leading tone* of $[Y]$ is the octave equivalence class of X 's 11th overtone. The *seventh tone belonging to* $[X]$, also called the *upper leading tone of* X 's *fifth overtone's equivalence class* is the octave equivalence class of the 7th overtone of Y . We also define these relations for the tones themselves: e.g. if $U \in [U]$ and $V \in [V]$ and $[U]$ is the leading tone of $[V]$, then we say that U is the leading tone of V .

We define the *perfect* X_1 *major scale* for a tone X_1 . Generally, a *seven-degree scale with base* X_1 is a set of tones $\{X_1, X_2, \dots, X_7\}$ where $f(X_i) > f(X_j) \Leftrightarrow i > j$ and $f(X_7) < 2f(X_1)$ (this is, every member of the scale is strictly less than one octave higher than the base). X_i is called the i th degree scale tone of the scale. According to the Hungarian notation, we denote the degrees and the operations among them with the elements of the prime field \mathbb{Z}_7 , but we use the capital Roman numeral for the integer $(n \pmod{7}) + 1$ instead of $n \in \mathbb{Z}_7$. The *perfect* X *major scale* is a seven degree scale with base X , where the frequency ratios of the neighbouring degree tones are respectively: $\frac{9}{8}, \frac{10}{9}, \frac{16}{15}, \frac{9}{8}, \frac{10}{9}, \frac{9}{8}, \frac{16}{15}$. where the last ratio is the ratio of the VIIth degree scale tone and the second overtone of X . For a perfect major scale, the following are approximately true, in the sense that the human ear cannot observe that they are false:

- (i) the degree VII tone is the leading tone of the degree I, III is the one of IV,
- (ii) IV is the upper leading tone of III, I is the one of VII,
- (iii) the interval between IV and I is a perfect fifth, the one of I and V is also,
- (iv) IV is the seventh tone belonging to I, and I the one belonging to V.

The sum of *twelve perfect fifths* starting from a tone X results a tone with frequency $\frac{531441}{4096} f(X)$, while the sum of *seven perfect octaves* gives a tone with frequency $128 f(X)$. The difference of these two tones is noticeable by an average person. However, if this deviation is equally spread along the whole interval, it locally cannot be perceived. Therefore one aims to fease the concept of *the circle of fifths*, i.e., 12 quasi-fifths equal to 7 octaves on a musical instrument each tone of which is a base of a seven-degree scale perceptually equivalent to a perfect major scale. The leading tone and upper leading tone/seventh tone connections between the quasi-perfect major scales could be used to make it possible to move from each major scale to the two with a base one quasi-perfect fifth higher respectively lower. This is the idea of the *equal-tempered piano*.

Definition 1. A countable set K of tones is an equal-tempered piano if

- (i) $A \in K$, where A is the normal a^1 tone with frequency 440 Hz,
- (ii) K has two tones X and Y the interval of which is at least 7 octaves,
- (iii) A tone X is an element of K if and only if $f(X) = (\sqrt[12]{2})^n f(A)$ for some $n \in \mathbb{Z}$ and $\exists Y, Z \in K$ such that $f(Y) < f(X) < f(Z)$.

According to this definition, A is the element of every – finite or infinite – equal-tempered piano. Thus, the octave equivalence classes of the piano's white keys (A, B, \dots, G) can be defined. Using a *well-tempered*, i.e., approximately equal-tempered piano that already actualized the circle of fifths, J.S. Bach showed that every tone of the equal-tempered piano can serve as a base of a quasi-perfect major scale, by composing his *Das wohltemperierte Klavier*, which contains one piece written in each major key of his well-tempered piano.

Enharmonic equivalence in the context of the 12-tone equal tempered scale means that two tones Y, Z originate from two different perfect major scales, but there is a tone X on the equal tempered piano from which neither Y nor Z is significantly different for the human ear. Enharmonic equivalence depends of the listener's own hearing and cultural background; here we follow the classicist European convention. Non-audible tones are called enharmonic if they have audible octave-equivalents that are enharmonic. The corresponding octave equivalence classes are also called enharmonic. If the tones A and B are enharmonic, we write $A \sim B$. It is easy to see that \sim is an equivalence relation. Further,

1. The interval of two neighbouring tones of the equal tempered piano is called *semitone*, the sum of two semitones (the distance of second neighbours) a *wholetone*. The sequence of tones 0, 2, 4, 5, 7, 9 and 11 semitones higher than an *arbitrary* piano tone is enharmonic to a perfect major scale.
2. If X and Y are two piano tones where Y is a wholetone higher than X , then the (only) piano tone Z such that $f(X) < f(Z) < f(Y)$ is the leading tone of Y and the upper leading tone of X , up to enharmonic equivalence.

The C major scale on the equal-tempered piano consists of the seven white keys. Moving *stepwise upwards* in the circle of fifths, one reaches the G, D, A, E, B major scales consecutively. At each step, one new tone appears in the scale, this is VIIth degree tone of the new scale. We denote this new tone as $X\sharp$, where X

is the element of the C major scale which has been replaced by the new, one semitone higher tone. This way the following tones appear consecutively: F \sharp , C \sharp , G \sharp , D \sharp , A \sharp , the leading tones to G, D, A, E, B respectively. After B, the next fifth step upwards leads to F \sharp . Now, starting from the C major scale again and move *stepwise downwards* fifth by fifth, we reach the F major scale first, which has exactly one scale tone that is outside the C major scale: instead of B, a one semitone lower tone occurs: the seventh tone with respect to F. Let X \flat denote the one semitone lower piano tone than the white key X, then moving downwards in the circle of fifths, we reach F, B \flat , E \flat , A \flat , D \flat and G \flat consecutively. \sharp and \flat marks can be multiplied. By construction, we have $\flat\sharp = \sharp\flat = \natural$ means the identity of the C major scale, and "multiplication" of \flat 's and \sharp 's is commutative.

Note that G \flat and F \sharp refer to the same (black) piano keys, these two tones are enharmonic, also D \flat is enharmonic to C \sharp etc. However, they are the same only under equal temperament: if we build a *perfect* A major scale, F \sharp is the VI degree scale tone there, while G \flat is reached if we move down in the D \flat major scale with 8 *perfect* fifth steps, and take the IV degree scale tone. It is a well-known experimental result that these actual G \flat and F \sharp are significantly different.

From this point, all major scales will be situated on an equal-tempered piano, with all degrees derivable from the C major scale with finitely many —in practice, usually 0, 1 or 2 — \flat s or \sharp s. The fifth-by-fifth sequence of sharpened scale tones of a major scale on the equal-tempered piano: F \sharp , C \sharp , . . . or the sequence of flattened scale tones of the major scale: B \flat , E \flat , . . . is called the major scale's *key signature*. Having established these scales on the piano, the traditional notation of musical intervals among their degrees can be established, see e.g. [2, Appendix E]. Also one can define *consonance* and *dissonance* of these intervals, cf. [2, Chapter 4].

By definition, an equal-tempered piano has to be *at least as wide as a real piano*, in order to make it possible that the piano covers 7 octaves (≈ 12 fifths). Let $d_2(X, Y)$ denote the interval of the notes X and Y of *any* equal-tempered piano K , measured in semitones. The construction of equal temperament implies the following, also if the equal-tempered piano is infinite.

Proposition 1. (K, d_2) is a metric space and d_2 generates the discrete topology.

Thus, if K is an equal-tempered piano, the Cartesian product K^n is also equipped with the discrete topology. The elements of K^n are called *chords*. Hence, we can speak about *Borel-measurable functions* $M : \mathbb{R}_0^+ \rightarrow K^n$, which we call *n-part pieces*. The k th voice of M is $\text{pr}_k \circ M$, where pr_k is the projection to the k th instance of the equal-tempered piano. We are interested in the *four-part case*; we define the special, homophonic four-part pieces in Section 3. There the voices (in increasing order of their numbers) are called, as conventionally, *bass*, *tenor*, *alto* and *soprano*. In classical harmony, each chord appearing in a homophonic four-part piece has to be a *triad* or a *seventh chord* in an correct four-part form.

Triad names are special elements of the factor space K^3 / \equiv on an arbitrary equal-tempered piano K , where \equiv is the octave equivalence relation. These contain scale tones or once altered tones from a certain seven-degree scale on K , and their main characteristic is that they consist of a k th, a $k+2$ th and a $k+4$ th

degree tone (mod 7) of the given major scale based on one of the twelve enharmonic equivalence classes of the equal-tempered piano. With this notation, we say that the triad is of degree k . There are four kinds of triad names for which we say that they *comply with classical harmony*, according to Table 1.

A *four-part version of a triad* – later in this article, simply: a *triad* – is an element of the piano power K^4 which consists of the tones of a triad name, exactly one of them in two voices. If the triad consists of the k th, $k + 2$ nd and $k + 4$ th degree scale tone of a seven-degree scale – these are called the *base*, the *third* and the *fifth* of the triad, respectively – on the equal-tempered piano, its *position* is determined by which tone it has in the bass. If in the bass there is the k degree tone, where k refers to the corresponding Roman numeral as before, then the triad is in *root position* (German–Hungarian notation of the triad: k), if the $k + 2$ degree tone is in the bass, then the triad is in *first inversion* (notation: k^6), and if the $k + 4$ degree tone, then in *second inversion* (notation: k_4^6).

Table 1. Triads (above) and seventh chords (below)

Name	Notation	$k \leftrightarrow k + 2$ interval	$k + 2 \leftrightarrow k + 4$ i.	$k + 4 \leftrightarrow k$ i.
Major triad	M	<i>major third</i>	minor third	perfect fifth
Minor triad	m	<i>minor third</i>	major third	perfect fifth
Diminished triad	d	minor third	minor third	<i>diminished fifth</i>
Augmented triad	A	major third	major third	<i>augmented fifth</i>

Name: (...) <i>seventh</i>	Third	Fifth	Seventh	Partial triads	Examples in major	Example in minor
augmented major	major	augmented	major	major, augm.	none	III
major minor	major	perfect	major	major, minor	I, IV	VI
major/dominant	major	perfect	minor	major, dimin.	V	V
harmonic minor	minor	perfect	major	minor, augm.	none	I
minor major	minor	perfect	minor	minor, major	II, III, VI	IV
semi-diminished	minor	diminished	minor	dimin., minor	VII	II
diminished	minor	diminished	diminished	dimin., dimin.	none	VII

Consider the union of a degree k and a degree $k + 2$ triad name on an arbitrary seven-degree scale. This is indeed an element of K^4 / \equiv , and it is called a *seventh chord name*. If $H \in K^4$ consists of the tones of a seventh chord name in any permutation of the voices, then H is called a *seventh chord*. This name comes from the fact that there is a seventh interval between the k and the $k + 6$ degree scale tones. The degree k and degree $k + 2$ triads are the *partial triads* of the seventh chord. The position of a degree k seventh chord inverson can be: (*root position*) *seventh chord* (German–Hungarian notation: k^7), *first inversion* (k_5^6), *second inversion* (k_3^4) and *third inversion* (k^2), if in the bass there is the k th, $k + 2$ nd, $k + 4$ th and $k + 6$ th degree tone of the seventh chord name, respectively. For the origin of these notations, we refer to [3, Book I., Section II.11].

The k th, $k + 2$ nd, $k + 4$ th and $k + 6$ th degree tones of a seventh chord are called *base*, *third*, *fifth* and *seventh* respectively. one is its *seventh*. If both partial

triads of a seventh chord name H comply with classical harmony, and there is no triad which is voicewise enharmonic to H , then we say that H *complies with classical harmony*. This second assumption is taken for excluding the *augmented triad* (see Table 1) from the set of seventh chord names, which can be represented as a seventh chord but is indeed just a triad. As a remark, we note that a root position *dominant seventh* (see Table 1) that complies with classical harmony may be *fifth deficient*, which means that it need not contain its fifth in any voice but instead the base in two voices (one of these voices is necessarily the bass).

Table 1 also shows the seventh chord types, with examples consisting of scale tones of the major and the minor key (see Section 2). For four-part versions of triads and seventh chords, we say that they *comply with classical harmony* if their name complies with classical harmony, in each of their voices the pitches (frequencies) accord to the conventional pitch interval of the voice, and —if the four voices are not of four different degrees— the duplication of tones is correct. For detailed duplication rules, we refer to the Hungarian version [8, Section 1.2].

2 Trichotomy of musical keys

After introducing the basic musical notions, we define keys based on the idea of key stability and functional tonality in classical harmony. We preliminarily ensure that our definition accepts the major and the minor keys, which have been used in Europe for five centuries, to be keys. Furthermore, our key definition gives us the possibility to find all possible key types. We prove that apart from major and minor there is exactly one more type.

Definition 2. *Let H be a seven-degree scale on the equal-tempered piano (consisting of scale tones and altered tones from the C major scale), with seven pairwise non-enharmonic scale tones. We say that H is a key if:*

- (i) *the V th degree seventh chord of H is dominant,*
- (ii) *all triad and seventh chord names that consist of the scale tones of H comply with classical harmony,*
- (iii) *if the k th degree seventh is dominant, then the degree $k + 3 \pmod{7}$ triad is major or minor $\pmod{7}$, with the $k + 3$ th degree scale tone one perfect fourth higher than the k th degree one.*

The condition (iii) means that the all dominant sevenths can *resolve to their tonic*, see Section 4. It follows that Definition 2 implies the next two properties:

Proposition 2. *In a key the first degree triad is major or minor, and the VII th degree scale tone is the leading tone of the 1st degree scale tone.*

The next lemma is a key observation of this section.

Lemma 1 (The Minor Lemma). *In any key H the following are equivalent:*

- (i) *the VI th degree scale tone is 8 semitones higher than the 1st degree one,*

- (ii) every type of seventh chord complying with classical harmony can be built from scale tones of H ,
- (iii) the VIIth degree seventh chord (built from scale tones) is diminished.

Proof. The definition of key implies that the sequence of intervals of the first degree scale tone and the other scale tones is: $(0, 2, ?, 5, 7, ?, 11)$, where the ?'s refer to unknown intervals. It is easy to see that the conditions of the lemma are equivalent to the condition that the sequence of intervals is $(0, 2, ?, 5, 7, 8, 11)$. The remaining ? stands for either 3 or 4, in order to satisfy the definition of key.

Proposition 3 (The trichotomy of keys). *Let X be an enharmonic equivalence class on an equal-tempered piano K . Then there are exactly three keys with first degree X , up to enharmonic equivalence. These are the major, the minor and the harmonic major (named by Rimsky-Korsakov in [6]) scales, with interval sequences $(0, 2, 4, 5, 7, 9, 11)$, $(0, 2, 3, 5, 7, 8, 11)$ and $(0, 2, 4, 5, 7, 8, 11)$, respectively. The latter two ones are the ones that satisfy the Minor Lemma.*

Proof. As in the previous proof, the key's definition implies that the interval sequence of an arbitrary key's scale is $(0, 2, ?, 5, 7, ?, 11)$. If the degree VI scale tone has sign 9, then the requirement that every triad and seventh chord built up from scale tones has to comply with classical harmony implies that the sign of the IIIrd degree tone is either 3 or 4. If this sign is 4, then the scale is the X -major scale. If the sign is 3, then the IVth degree seventh chord built up from scale tones is a dominant seventh, but the interval between the IVth and the VIIth degree scale tone is not a perfect fourth but a tritone (enharmonic with 6 semitones/3 wholetones). Therefore in this case we do not obtain a key.

From the 1500s, European music is determined by the major-minor duality. The harmonic major key differs by only one scale tone (degree VI) from the major scale and also by only one scale tone (degree III) from the minor one, and therefore the listener automatically tries to perceive it as minor or major. Whether a harmonic major piece gives the feeling of minor or major depends also on which degree chords the piece uses, cf. [8, Section 1.8].

The *key signature of a harmonic minor or harmonic major key* is the key signature of the major scale which has the degree I scale tone of the minor key as degree VI scale tone. These major and minor scales are called *parallel*.

3 Topology of the homophonic four-part setting

In this section, we present a continuous time model of classical harmony, given by *homophonic four-part pieces*. Our notions allow for some non-feasible musical phenomena, such as infinite pieces and chords accumulating in one point in time (referred to as *packing point*). This way, one can also handle *periodic pieces* without ending in finite time, which is often aimed in both classical and popular music, cf. [8, Section 2.4]. The continuous approach allows us to define the genre of *Bach's chorale harmonizations* mathematically precisely, which is not possible if one uses only *chord sequences* to describe homophonic four-part pieces.

Definition 3. Let K be an equal-tempered piano. $M : \mathbb{R}_0^+ \rightarrow K^4$ be a four-part piece (see Section 1). M is called a homophonic four-part piece if:

- (i) Each element of $\text{Ran } M$ is a four-part version of a triad or a seventh chord (in some inversion) that complies with classical harmony,
- (ii) each voice of each element of $\text{Ran } M$ only contains tones that can be derived from the C major scale on K using the system of \flat 's and \sharp 's,
- (iii) for all $H \in \text{Ran } M$, we have that $M^{-1}(H) = \{x \in \text{Dom } M \mid M(x) = H\}$ is a disjoint union of intervals closed on the left and open on the right.

Definition 4. If M is a homophonic four-part piece, $B(M)$, the smallest (left-closed, right-open) interval that contains $\text{Dom } M$ is called the cover of M .

As we mentioned in the introduction, homophony means that if in a point in time one voice starts to play a new tone, all other voices do so. It follows that if M is a homophonic four-part piece, then if there is a pause at time $t \in B(M)$ in at least one voice of M (i.e., $t \notin \text{Dom } pr_i \circ M$), then this is actually a *general pause*, i.e. pause in all voices. Also, one can prove that the connected components of pauses are also intervals closed on the left and open on the right.

Using the point (iii) of Definition 3, it is easy to see that any homophonic four-part piece M is *right-continuous*. According to the fact that M takes values in a discrete space, this implies $\forall t_0 \in \text{Dom } M \exists \delta > 0 : \forall t \in [t_0, t_0 + \delta[\ M(t) = M(t_0)$. Now we define some special points of homophonic four-part pieces.

Definition 5. Let M be a homophonic four-part piece.

- (i) $t = \inf \text{Dom } M$ is the starting point of M ,
- (ii) $t = \sup \text{Dom } M$ is the endpoint of M ,
- (iii) $t \in \text{Dom } M$ is a chord-changing point of M if $\exists \varepsilon > 0, \exists H_1 \neq H_2 \in K^4$ such that $\forall x \in [t - \varepsilon, t[\ M(x) = H_1$ and $\forall x \in [t, t + \varepsilon[\ M(x) = H_2$. Let $A(M)$ denote the set of the chord-changing points of M .

In the following, \vee means logical "or" and \wedge means logical "and".

Definition 6. A homophonic four-part piece M has infimum of chord lengths defined as $\inf_{t \in \mathcal{D}(M)} \sup \{r_1 + r_2 \mid r_1, r_2 \geq 0 \wedge \forall x \in [t - r_1, t + r_2[: M(x) = M(t)\}$.

The proof of the next proposition is left for the reader.

Proposition 4. Let M be a homophonic four-part piece. Then $A(M)$ is countable. If the infimum of the chord lengths of M is positive, then $A(M)$ has no accumulation point, which implies that $A(M)$ is finite if $\text{Dom } M$ is bounded.

The following definition accounts for a non-feasible musical effect coming from our topological model, which is playing infinite music in finite time.

Definition 7. Let M be a homophonic four-part piece. $t \in \overline{\text{Dom } M}$ is a packing point of M if $\forall \varepsilon > 0 \ [t - \varepsilon, t[$ contains infinitely many chord-changing points or isolated boundary points of $\text{Dom } M$.

Packing points can have interesting applications in the spirit of [4, part I.; Decision]. But in the usual model in classical harmony, packing points do not occur, and neither do pieces without packing points but with infinitely many chords.

Definition 8. *A homophonic four-part piece M is feasible if*

- (i) *the infimum of chord lengths of M is positive, and if $\text{Dom } M \neq B(M)$, then the infimum of general pause interval lengths is also positive,*
- (ii) *and $\overline{\text{Dom } M}$ is compact.*

We have finished the construction of *natural parametrization* of homophonic four-part pieces. *Playing functions* describe the performance of these pieces with non-constant velocity.

Definition 9. *Let M be an n -part piece for a certain $n \in \mathbb{N}^+$ and $\theta : [0, \infty[\rightarrow [0, \infty[$ a continuous, strictly increasing function, for which $[0, \infty[$ can be divided into countably many disjoint intervals $(I_i)_{i \in \mathbb{N}}$ joining each other and altogether covering $[0, \infty[$, such that restricted to the interior of each interval I_i , θ is twice continuously differentiable, θ' nowhere vanishes and $\inf_{n \in \mathbb{N}} \lambda(I_n) > 0$.*

Then θ is called a playing function. The name of $M \circ \theta|_{B(M)}$ is the playing of M that belongs to θ . For $t \in B(M)$, $\theta'(t)$ is called the playing velocity and $\theta''(t)$ the playing acceleration in the point t , if they exist. $\theta \equiv 1$ gives the naturally parameterized n -part piece. The set of playing functions is denoted as $PL(\mathbb{R}^+)$.

It is easy to verify that on bounded intervals playing functions are absolutely continuous. The following definition is based on this.

Definition 10. *If $\theta \in PL(\mathbb{R})$, M is a homophonic four-part piece and A is a Lebesgue-measurable subset of $B(M)$, then the length of the part A of piece M by the playing function θ is $\mu_\theta(A) = \int_A 1\theta(dx) = \int_A \theta'(x)dx$.*

For the proof of the next proposition, see [8, Section 1.4].

Proposition 5. *$PL(\mathbb{R})$ is a group under the composition of playing functions.*

Now we turn our attention to the mathematical definition of the genre of Bach's chorales. We emphasize that chorale is an actual musical genre from Baroque, and hence its characteristics are originally non-prescriptive. Therefore however precisely we define a chorale, our definition may only be correct for the majority of the pieces, with certain exceptions.

Definition 11. *Let M be a homophonic four-part piece with $x \in \text{Dom } M$ and $M(x) = H$. Then the area of $M(x)$ is the connected component of $M^{-1}(H)$ containing x .*

The halving of $I = [a, b[\subseteq B(M)$ in the playing belonging to $\theta \in PL(\mathbb{R})$ is dividing I into two disjoint intervals closed on the left and open on the right I_1, I_2 which together cover I and $\mu_\theta(I_1) = \mu_\theta(I_2)$.

Using this, our definition for four-part chorale is the following.

Definition 12. A four-part piece \mathfrak{K} is a four-part chorale if there is a feasible, naturally parameterized, pauseless homophonic four-part piece M such that $\exists c > 0: \forall x \in \text{Dom}(M)$ the length of the area of x by the identic playing function is c , and

(1) \mathfrak{K} can be derived from M with using the following steps, the so-called figurations. They are used for a finite number of $x \in \text{Dom } M$ and the figurations excluding each other are not done at the same time.

Types of the figurations are:

Chord duplication Halve the area of $M(x)$ by the identic playing function (natural parametrization), and in the first half of the area keep $M(x)$ for $\mathfrak{K}(x)$, in the other half $\mathfrak{K}(x)$ is one constant triad or seventh chord different from $M(x)$.

Suspension Halve the area of $M(x)$ by natural parametrization, in the second half of the area keep $M(x)$, in the first half, in one or two voices change the appropriate tone of $M(x)$ one step higher and keep the remaining voices.

Advancement Halve the area of $M(x)$ by natural parametrization, in the first half of the area keep $M(x)$, in the second half, in exactly one voice write a step higher or lower tone, which is equal to the tone in the same voice of the next chord after $M(x)$.

Accented passing tone Suppose that there is a third skip in some voice(s) of M at arriving at or departing from $M(x)$. Then halve the area of $M(x)$, and on the half which is closer to the interval of the neighbouring chord that is involved in the third skip, instead of $M(x)$, write a tone the degree of which is between these two tones' degree. Note that only one accented passing tone per one chord area of M is accepted.

(2) The given four-part piece \mathfrak{K} is meant to be associated with a canonic playing function θ that differs from the identic playing in the following: $\exists m \in \mathbb{N}$ such that θ changes the length of every m th chord interval of M to k times greater than originally, where $k \in]1, 2[$ is a conventionally accepted factor. Then we say that there is a pause on every m th metric unit.

4 Convergence area of a key. Functions and tonality

Let T be a key with degree I scale tone X in an enharmonic equivalence class on the equal-tempered piano K . The *convergence area* of T is defined $CA(T) = \{(G_i, L_i) | i = 1, \dots, N\}$, where $N \in \mathbb{N}$ and $\forall i$, G_i is a fixed triad or seventh chord name of K in a certain inversion and L_i is the list of the accepted four-part versions of G_i , according to the definitions of compliance with classical harmony from Section 1. In simplified notation, we also call G_i an element of the convergence area, and view $CA(T)$ as the set of chords belonging to T .

We explain the meaning of convergence area of T as follows. Roughly speaking, a chord X that complies with classical harmony and is deducable from the C major scale on K by \sharp s and \flat s is considered to be an element of $CA(T)$ if:

1. it is built from the scale tones of T and regularly used in homophonic four-part pieces associated to this key. Some inversions of some chords are excluded for their too strong dissonances, e.g. diminished triads may only stand in first inversion, moreover in the case of degree VII triads, the duplicated tone of the triad must be the third. Apart from I_4^6 and IV_4^6 , triads in third inversions are not used. All such seventh chords, which are called *diatonic seventh chords*, are used in all inversions, though some of them very rarely.
2. if $X \in CA(T)$ has a tone outside scale of T , then X is called an *altered chord* of T . Such X is convergent if and only if it can lead to chords in $CA(T)$ built from scale tones, without violating any chord-changing compositional principles, in such a way that was usual in the practice of Viennese classical composers. A full list of convergent chords in minor and major key was given in [8, Appendix D], an explanation of altered chords in [8, Section 1.7].

After introducing convergence areas, we define weak tonality.

Definition 13. *Let M be a homophonic four-part piece and t an accumulation point of $\text{Dom } M$. We say that M is weakly tonal in the point t with key T if there is a connected open neighbourhood U of t such that $\forall x \in (U \setminus \{t\}) \cap \text{Dom } M$, $M(x)$ is the element of $CA(T)$.*

Weak tonality is sufficient in the case when there are no modulations among different keys, but classical harmony has stronger measures on key stability, especially for establishing new keys after modulations. This involves the notion of musical *functions*: the *tonic*, *dominant* and *subdominant*. In the following, the k th degree triad or seventh chord of the key T of a tone will mean the one built from scale tones of T . The *leading tone/seventh tone* of a key T will refer to the leading tone/seventh tone of the key's 1st degree scale tone. The *leading tone* of a *diminished triad* or *diminished seventh* is, by definition, its base.

Definition 14. *Let T be a key, $H \in CA(T)$ be a major or diminished chord, i.e. major triad, diminished triad, major (dominant) seventh or diminished seventh, and $G \in CA(T)$ a major or minor third. We say that H resolves to G if*

- (i) H contains the leading tone of (the base of) G , and
- (ii) if there is a tone x belonging to H that is not a scale tone in the major key built on the base of G , then x is the upper leading tone of the fifth of G .

Definition 15 (Dominant function (D) and secondary dominant property.). *$X \in CA(T)$ has the dominant function in the key T if it resolves to the first degree triad of T . $Y \in CA(T)$ is a secondary dominant chord if it resolves to any other major or minor chord built from the scale tones of T .*

Definition 16 (Tonic function (T)). *$X \in CA(T)$ has the tonic function in the key T if*

- (i) X contains a 1st and 3rd degree tone of T , the first one from the scale T ,
- (ii) if X contains the leading tone of T , then it is the seventh tone of X ,
- (iii) if X is secondary dominant, then X is a 1st degree major triad,

(iv) X has no augmented and no diminished partial triad.

Definition 17 (Subdominant function (S)). $X \in CA(T)$ has the subdominant function in the key T , if

- (i) X contains the IVth and VIth degree scale tone of T , possibly both altered,
- (ii) if X is secondary dominant (itself, not just up to enharmonic equivalence), then it resolves to the Vth degree triad. Moreover, the VIth degree tone of X has neither more \sharp s nor more sharpening \flat s than the key signature of T .
- (iii) The set of tones of X and the one of the Ist degree seventh chord of T have no other common tone than the Ist degree scale tone. Moreover, X contains no altered Ist degree tone.

In Table 2, we present the most typical tonic, dominant and subdominant chords built of scale tones. Here, if a root position triad participates in the table, then its first inversion has the same function. A seventh chord participating in the list has the same function as any of its inversions. Note that we require that the *seventh degree diminished seventh chord*, is also the element of $CA(T)$ if T is a major key, where it is an altered chord. The Minor Lemma (Lemma 1) guarantees that this chord is built from scale tones in a minor or harmonic major key. The non-altered convergent chords which are not listed in the table have no certain function. E.g., this applies for the IIIrd degree triad, as it is considered to be pending between tonic and dominant function, and many diatonic seventh chords also do not have a certain function.

Table 2. Chords belonging to the three functions in the three different kinds of keys.

Type of T	Tonic chords	Dominant chords	Subdominant chords
major	I, VI, VI ⁷ , I ⁷	V, VII ⁶ , V ⁷	II, IV, II ⁷
minor	I, VI, VI ⁷	V [#] , VII ^{6#} , V ^{7#} , VII ⁷	II, IV, II ⁷
harmonic major	I [#] , VI ^{5#} , VI ⁷ _{5#}	V [#] , VII ^{6#} , V ⁷ , VII ⁷ _#	II, IV, II ⁷

The degree I triad is called the *tonic main triad* of the key T , the degree IV one is the *subdominant main triad* and the degree V one is the *dominant main triad*. *Authentic step* means two different things in classical harmony. On the one hand, modulation (key change) to the one fifth higher key (the *dominant* key), without changing the type of key. Among triads this means a V→I or I→IV type chord progression. On the other hand, function change $D \rightarrow T$ in a certain key in general. Similarly, *plagal step* means two things. On the one hand modulation to the one fifth lower —the *subdominant*— key, and among triads making a I→V or IV→I step, on the other hand function change $T \rightarrow D$ in certain key.

A *cadence* is a chord progression consisting of at least two chords that is considered to be appropriate for finishing a piece. In view of this, we can call the dominant→tonic or tonic→subdominant steps *authentic cadences* and the tonic→dominant and subdominant→tonic steps *plagal cadences* in certain cases.

In a given key, a *complete authentic cadence* is a chord progression with $T \rightarrow S \rightarrow D \rightarrow T$ function sequence, while a *complete plagal cadence* is a chord progression with $T \rightarrow D \rightarrow S \rightarrow T$. It is a well-known fact that complete authentic cadences are the most applicable for finishing a piece, for which there are many arguments, but it is hard to get a full explanation, cf. [2, Section 5.11]. Most of the classical, romantic and also recent popular music is based on $D \rightarrow T$ resolutions, supported by complete authentic cadences using the function S .

In the following, we present strong, *functional tonality*. For this, we need to provide our first axiom of classical harmony, in particular about *modulations*.

Definition 18 (Modulation). *If there are keys T_1 and T_2 for the homophonic four-part piece M such that $\text{Dom } M$ has a subset $Z = [a, b[$, for which $M|_Z$ is feasible, and $\exists r_1 > 0, r_2 > 0$ such that on the whole set $B_{r_1}(a) \cap \text{Dom } M \setminus Z$, M is weakly tonal with key T_1 and on the whole set $B_{r_2}(b) \cap \text{Dom } M \setminus Z$, M is weakly tonal with key T_2 , then $\forall W \subseteq Z$ we say that W belongs to a $T_1 \rightarrow T_2$ modulation. We also say that M modulates on Z from T_1 to T_2 .*

Thus, we demand that modulations themselves be *feasible* and *pauseless*: they need to last until a finite time, without general pauses and packing points.

Axiom 1 (First modulational axiom). Let M be a homophonic four-part piece. If M complies with classical harmony and contains a $T_1 \rightarrow T_2$ modulation, then $\exists [a, b[\subseteq \text{Dom } M$ such that $M(a)$ is the degree I triad of T_1 (built from scale tones), $M(b-) = \lim_{x \rightarrow b-0}$ is the degree I triad of T_2 (also consisting of scale tones), and M is weakly tonal in a with key T_1 , M is weakly tonal in b with key T_2 , and $[a, b[$ is the largest interval which belongs to this $T_1 \rightarrow T_2$ modulation.

Now we have all notions that we need for defining functional tonality. The idea for tonality of a piece in one point is to assign a key T to the point as a limit, requiring that all three functions of T occur in the vicinity of the point.

Definition 19 (Local tonality with a given key.). *Let M be a (not by all means homophonic) four-part piece and t an accumulation point of $\text{Dom } M$. M is tonal in the point t with key T if there is a connected open neighbourhood U of t such that $V = (U \setminus \{t\}) \cap \text{Dom } M$, $M(x)$ satisfies the following conditions:*

- (i) M is weakly tonal with key T in every point of V ,
- (ii) $M[V] = \{M(x) \mid x \in V\}$ contains at least one chord from all functions of T ,
- (iii) if $t \notin \text{Int } \text{Dom } M$, then only triad-valued points of $\text{Dom } M$ accumulate to t .

Definition 20 (Local tonality via modulation.). *Let M be a homophonic four-part piece, t an accumulation point of $\text{Dom } M$ and $T_1 \neq T_2$ two keys. M is tonal in t and modulates from T_1 to T_2 if there is a connected open neighbourhood U of t such that $\exists [a, b[= V \supseteq U$, where V belongs to a modulation (see Definition 18), which complies with classical harmony apart from the chord-changing points.*

Definition 21 (Tonal piece). *Let M be a homophonic four-part piece, $A \subseteq \text{Dom } M$. M is tonal on A if $\forall x \in \overline{A}$, M is tonal in x by Definition 19 or 20.*

5 Axioms and the fundamental theorem of tonality

The most well-known classical compositional principles are the *chord-changing or voice-leading rules*. The goal of the homophonic four-part model is to describe the kind of chord progression and voice-leading between chords that classical harmony accepts. Often the formal rules of classical harmony do not tell how to write pieces but what to avoid: it forbids some kinds of chord progressions (e.g. V→IV steps in some cases) and some kinds of voice-leading (e.g. parallel octaves or augmented second steps). Virtually, this property of the axiom system gives the freedom to actually write *pieces of art* and not just ‘correct examples’ complying with classical harmony. What one *should* write follows from the practice of Viennese classicist authors, cf. [2, Section 5.11] or [8, Section 1.6.3].

Now, we present our axiom system, which consists of:

- the rules of compliance of triads and seventh chords with classical harmony,
- the definition of correctable piece (global level of the pieces),
- the compositional principles for modulations (semi-global level, describing global properties of a modulational segment of a piece). These are summarized in Section 6, see also [5, p. 36],
- the compositional principles for chord-changes (local level).

Our main result, the *fundamental theorem of tonality* helps us embed the chord-changing rules in a mathematical axiom system for classical harmony. Its equivalent condition for tonality gives a general framework according to which pieces can comply with classical harmony *apart from the chord changes*.

Definition 22 (Correctable piece). *Let M be a homophonic four-part piece, $a \in \text{Dom } M$, $b \in \overline{\text{Dom } M} \cup \{\infty\}$, $N = [a, b[\subseteq B(M)$. $M|_N$ is called a correctable piece if all the following conditions are satisfied*

- (i) $M|_N$ has a positive infimum of chord lengths,
- (ii) N is the disjoint union of a finite odd number $2n + 1$ ($n \geq 0$) of left-closed, right-open intervals (I_1, \dots, I_{2n+1}) such that $\forall i \in \{1, 2, \dots, 2n + 1\}$, $I_i \cap \text{Dom } M \neq \emptyset$. Further, $\forall 0 < k \leq n$ in the whole interval I_{2k} , M modulates complying with classical harmony apart from the chord-changing points, and $\forall 0 \leq k \leq n$ for $I_{2k+1} \cap \text{Dom } M$ there is a unique key T_k such that $\forall G \in M[I_{2k+1}] = \{H \in K^4 | \exists x \in I_{2k+1} : M(x) = H\}$, $G \in CA(T_k)$, and $M[I_{2k+1}]$ contains at least one tonic, one dominant and one subdominant chord of T_k ,
- (iii) If $x \in N \cap \partial M$, then only triad-valued points accumulate to x .

Moreover, if $t \in \text{Int Dom } M \cap N$, then if there is no chord change in t forbidden by the axioms regarding chord changes (see below), then we say that M complies with classical harmony in t and N is a classical neighbourhood of t .

Finally, if $t_0 \in M$ is such that $t_0 \in \text{Dom } M$ but $\exists \varepsilon > 0$ such that t_0 is the starting point of $M|_{N \cap [t_0 - \varepsilon, \infty)}$, then according to the definition of strict four-part setting, $\lim_{t \rightarrow t_0 + 0} M(t) = M(t_0)$. Then, if $\exists r > 0$ such that $\forall x \in]t_0, t_0 + r[: x \in N$ and M complies with classical harmony in x , then we say that M complies with classical harmony in t_0 . If $t_0 \neq a$, then we call N a classical neighbourhood of t_0 .

Definition 23. *If a homophonic four-part piece M complies with classical harmony in x , $\forall x \in \text{Dom } M$, then we say that M complies with classical harmony.*

Theorem 1 (The fundamental theorem of tonality). *Let M be a homophonic four-part piece that is pauseless ($\text{Dom } M = B(M)$) and feasible. Then M is tonal (on $\overline{\text{Dom } M}$) if and only if it is correctable, i.e. if it complies with classical harmony (on $\text{Dom } M$) apart from its chord-changing points.*

Proof. The fact that the condition of the theorem is sufficient for the tonality is almost clear from Definitions 19 and 22, therefore we omit this part of the proof.

We show that the condition is necessary for the tonality. Let M be tonal, feasible and pauseless. For all $t \in \overline{\text{Dom } M}$, let U_t be an open neighbourhood of t that shows its tonality. If possible, let us choose U_t such that it shows key and not modulation. Since $\text{Dom } M$ is a bounded subset of \mathbb{R} , it can be assumed that $\forall t \in \overline{\text{Dom } M}$, U_t is a bounded open interval. Then, $\bigcup_{t \in \overline{\text{Dom } M}} U_t$ is an open cover of $\overline{\text{Dom } M}$. Since $\overline{\text{Dom } M}$ is compact, this has an open subcover, which we denote by $\bigcup_{i=1}^n U_i$. Without loss of generality, we can assume that $U_i =]a_i, b_i[$, where $a_i < a_j \Leftrightarrow i < j$ and $b_i < b_j \Leftrightarrow i < j$, moreover that $\inf U_1$ is the starting point of M and $\sup U_n$ is the endpoint of M . The tonality of M guarantees that in U_1 , M is tonal with some key T_1 . Let us start a sequential process with $V = U_1$, $\mathfrak{I} = \emptyset$, $j = 1$ and $T = T_1$ in order to divide $B(M)$ into an interval system that shows that M is correctable.

1. If $\forall k > j$ we have that in U_k , M is tonal with same key as in V (or $\sup V$ is the endpoint of M), then let us append $(V \cup \bigcup_{k>j} U_k) \cap \text{Dom } M$ to \mathfrak{I} , as the last interval for showing correctability. In this interval M is tonal with key T . Also, only triad-valued points of $\text{Dom } M$ accumulate to the only two boundary points of M , which are the starting point and the endpoint.

2. Else, $\exists k > j$ such that in U_k there is a key T' , since Definition 19 implies that in $\sup \text{Dom } M$ there has to be tonality with a key. Then let us define $s = \sup\{x \in]\inf V, \sup U_k[\mid M \text{ is tonal in } x \text{ with key } T\}$ and $i = \inf\{x \in]\inf V, \sup U_k[\mid M \text{ is tonal in } x \text{ with key } T'\}$. Note that $\sup U_k$ is finite, and the tonality of M implies weak tonality with key T in s and weak tonality with key T' in i , therefore $s \leq i$. By Definition 1, $s < i$ follows. Then, on the entire interval $[s, i]$, M modulates from T to T' complying with classical harmony apart from the chord-changing points. Let us append $[\inf V, s[$ (as an interval with key T) and $[s, \sup U_k[$ (as an interval of a $T \rightarrow T'$ modulation) to the set \mathfrak{I} of the intervals showing the correctability of M . Let us put $T = T'$, $j = k$ and $V = [i, \sup U_k[$ and return to the starting alternative of the sequential process.

Each time the process restarts, the endpoint of the current V is the endpoint of U_k for a k larger by at least 1 than the one in the previous turn. This ensures that the process is finite, further the number of turns is not more than n : when $\sup V = \sup \text{Dom } M$ holds, the process is finished. The intervals given by the process show that M is correctable: the intervals with an odd index are intervals with key and the ones with an even index contain modulation from the previous interval's key to the following one's. Thus, each point of $\text{Dom } M \setminus A(M)$ has a classical neighbourhood containing $B(M)$. This finishes the proof. \square

the tonal piece M complies with classical harmony apart from chord-changing points:

- A tonal piece may have a packing point.
- A tonal piece M with $B(M) = \mathbb{R}_0^+$ may have no packing point but the infimum of the lengths of chord intervals can be still zero (in this case the sum of the chord lengths must be infinite).
- A tonal piece M with $B(M) = \mathbb{R}_0^+$ and positive infimum of chord lengths can contain infinitely many modulation intervals. In this case it can occur that $\forall t \in]\inf \text{Dom } M, \infty[$, we have that $M|_{[0,t]}$ is correctable but M itself is not correctable. In this case, a *final key* of M cannot be defined. The final key is a main characteristic of finite feasible tonal pieces, in particular, in classicist music the names of the pieces of music often contain the final key of the piece (e.g. symphony in G major etc.).
- A tonal, feasible piece is not by all means correctable if it is not pauseless. Indeed, take a piece that complies with classical harmony and contains no modulations, write some positive amount of it, and then continue with the same piece in a different key. The two connected components of the resulting piece with classical harmony by themselves comply, but the entire piece is not correctable because the modulation between the two keys is missing. For an example for such a piece, see Figure 5.

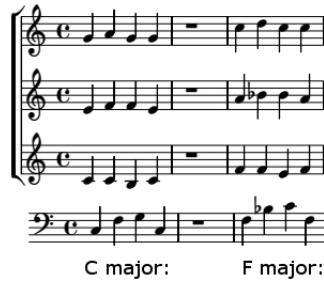


Fig. 1. A non-pauseless tonal piece that is not correctable due to lack of modulation.

Knowing the fundamental theorem of tonality, we can embed the basic chord-changing compositional principles of classical harmony (see e.g. [3, p. 30–183.] or [2, Section 5.11]) in our mathematical axiom system. We present a scheme how these compositional principles can be stated, according to Theorem 1.

Axiom 2 (Scheme of chord-changing rules). Let M be a tonal, feasible, pauseless homophonic four-part piece and t a chord-changing point of M . If M complies with classical harmony in t , then [*conditions on the chord change in t*].

This scheme guarantees that new chord-changing compositional principles can be added to the axiom system of classical harmony as long as the compositional

principles do not contradict each other. We note that one of the chord-changing rules, the *principle of least motion*, is very hard to formalize mathematically precisely in full generality. but there exist interesting mathematical results about this principle in the literature, see e.g. [7, p. 4–6.].

6 Modulations

In this last section, we sketch the classical compositional principles for modulations with seven chords. Here we do not enumerate all exact details of the technically rather complicated compositional principles themselves, neither the *altered chords*, the elements of the convergence areas of the keys not detailed so far. Our whole model for modulations that can be found in my Bachelor’s thesis at [8, p. 55–59.], in which all the altered chords of the keys and precise formulations of the modulational axioms takes place, is mathematically complete—but still far from universal, as it only describes modulations consisting of seven chords. The seven chords of the modulation do not have to ensure that there is tonality with key T_1 in the starting point of the modulations, but the new key T_2 has to be established by a complete authentic cadence, according to the compositional principles.

Definition 18 for modulations guarantees that in the context of modulations it is enough to consider feasible and pauseless homophonic four-part pieces. Pauselessness ensures that chords that cannot follow each other by actual chord-changing also will not occur directly after each other, separated by a pause. In general, pauses can weaken the impact of irregular chord progression and there are some examples in music history when composers use this. But in the case of modulations, in Viennese classicism, the basic aim is to make the key change as smooth as possible and to find some connection between the beginning key and the target key, therefore such trickery is not advised.

In the following, we establish the notions that are necessary to state the remaining compositional principles for modulations. First, we define pauseless extensions of general homophonic four-part pieces, in order to obtain a completely pauseless paradigm for the modulations that incorporates non-pauseless pieces as well. Using these, we define chord sequences, which provides a simpler interpretation for feasible pieces than the one described in Section 3. In the same time, note that the chorales cannot be defined mathematically precisely using only chord sequences, therefore also the continuous time construction from Section 3 is useful.

Let M be a homophonic four-part piece that has no packing point apart from its endpoint, then we have a (possibly finite) sequence of disjoint consecutive intervals (I_i) of which $B(M)$ consists, for all of which either $I_i \subseteq \text{Dom } M$ and M has the constant value of a chord with $\inf I_i$ and $\sup I_i$ being either chord-changing points or boundary points of $\text{Dom } M$, or I_i is a maximal general pause interval in the sense that $\inf I_i$ and $\sup I_i$ are boundary points of $\text{Dom } M$. In this case, there is a simple way to construct a pauseless extension \bar{M} of M , given by an extension from $\text{Dom } M$ to $B(M)$, this is called the *right-invariant pauseless*

extension of M :

$$\overline{M}(t) = \begin{cases} M(t), & \text{if } t \in \text{Dom } M, \\ M[I_i] = M(\sup\{u \in \text{Dom } M \mid u < t\}), & \text{if } t \in I_{i+1}, I_{i+1} \cap \text{Dom } M = \emptyset. \end{cases}$$

In this case, $(\overline{M}(t_i), t_i \in A(\overline{M}))$ is called the *chord sequence* of M , here t_i 's follow each other in their order in \mathbb{R}_0^+ . We omit the proof of the following proposition, which shows the role of the playing function group $PL(\mathbb{R})$ in the topology of four-part pieces.

Proposition 6. *Let M_1 and M_2 be homophonic four-part pieces such that the values of the chord sequences of M_1 and M_2 are the same. Then $\exists \theta \in PL(\mathbb{R})$: $M_2 = M_1 \circ \theta$.*

When describing modulations, we will not make any difference between feasible homophonic four-part pieces which have the same chord sequence. The expressions 'a chord sequence is tonal/is correctable/complies with classical harmony' will be used in the sense that every homophonic four-part piece with the given chord sequence has this property.

Firstly, we have to establish a connection between modulational axioms and the definition of the correctable piece. We call modulations which satisfy not only Definition 18 but also the first modulational axiom (Compositional principle 1) *basic modulations*. For a $T_1 \rightarrow T_2$ modulation that complies with classical harmony apart from chord changes, the next necessary condition that we require is tonality in the starting point of the first degree triad of T_1 that opens the modulation, with key T_1 , and to also tonality in the endpoint of the first degree triad of T_2 that opens the modulation (the existence of these chords is guaranteed by the first modulational axiom).

Now we can turn to the basic idea of modulations complying with classical harmony: the chord sequence of the —pauseless, feasible— $T_1 \rightarrow T_2$ modulation section has to be able to be divided into three disjoint segments (left-closed, right-open intervals) that cover the whole chord sequence [5, p. 36]:

Neutral phase (N) In this segment, which is opened by the 1st degree triad, the key of the piece is still T_1 (i.e., each element of N is a member of $CA(T_1)$), but there are no secondary dominant chords. In the whole modulation after the first chord there is neither in T_1 nor in T_2 any root position 1st degree triad until the tonic main triad of T_2 occurs and closes the modulation.

Fundamental step (F) If T_1 and T_2 are of the same type and they are neighbours in the circle of fifths, this whole segment may be empty. Otherwise here dominant chords of different keys follow each other. Only the last can be a (major/diminished/in the case of minor T_1 and minor T_2 augmented) triad, the ones before have to be seventh chord inversions. These seventh chords have to follow each other by *elision*¹. The last chord of F may be a triad and the previous chord may resolve to it.

¹ We generally use the word 'elision' for chord progression of inversions of different seventh chords, without the first seventh chord resolving to its tonic. The lack of

Cadence (C) The modulation has to be finished by a complete authentic cadence in the new key T_2 , this shows and stabilizes the tonality in the new key. It may occur that we do not write a cadence in each key, but a modulation progress is only finished when we reach a cadence in some key. It is sure that the last chord before the closing degree I triad is the degree V triad or degree V dominant seventh chord of T_2 in the segment C . The tonic and subdominant chords of T_2 that precede this chord belong already to C (and not F) if and only if they are built up from scale tones in T_2 , otherwise they belong to F .

If a chord sequence of a basic modulation has all the properties that we have introduced in this section, and it is the member of one of the following three modulation types, then we say that it complies with classical harmony apart from its chord-changing points. If its chord-changings are also correct, we say that the modulation complies with classical harmony. The three possible modulation types are:

Diatonic For the last chord H of N we have $H \in CA(T_2)$, and every chord after this is convergent in T_2 . This time F usually consists of at most one chord. This is the smoothest possible key change type, but it is often not possible between keys further away from each other.

Enharmonic The last chord of N or the first chord of F is an element of T_1 that is enharmonic with some element of $CA(T_2)$. The most common enharmonic modulation types use the enharmonic equivalence of diminished seventh chords or augmented triads in different keys. After this chord occurs, we consider it as an element of $CA(T_2)$, and make a chord progression in T_2 ending with an authentic cadence.

Chromatic There is elision in the modulation chord sequence. Very far away leading modulations, such as C major \rightarrow F \sharp major can be most conveniently feased this way. In most of the chromatic modulations $\sharp F \geq 2$ holds.

These three categories do not exclude each other pairwise, while it is difficult to accomplish a modulation that is both diatonic and enharmonic at the same time. In the music score collection of the thesis we show examples of both enharmonic and chromatic and both diatonic and chromatic modulations.

Modulational compositional principles finish our axiomatization work. By providing more and more detailed chord-changing compositional principles based on the research of real classical music pieces and Bach's chorales, the four-part model can be refined again and again. The logical ordering of musical notions and the mathematically simpler results in this paper can now be used for writing

resolution is expressed by 'elision', the greek word for 'omission'. Chord progression using elision always has to use *chromatics* in order to make it possible to comply with classical harmony. Chromatics means the sequence of at least two *semitone steps* after each other in one voice. This semitone sequence has to be such that for each segment of it where the consecutive semitone steps are taken in the same direction, there exists a key in the scale of which, exactly every other step changes degree (i.e., every other step is a minor second and the remaining steps are augmented primes).

a new classical harmony coursebook. We also plan to do experiments on the possibilities and barriers of composing four-part chorales by Markov models, revising the results of [9].

Acknowledgements I would like to thank my Bachelor’s thesis supervisor Ákos G. Horváth for his support, corrections and advice for my research. I would like to thank Robert S. Sturman and Scott McLaughlin for discussing with me about mathematical music theory at the University of Leeds (UK). I would like to thank Máté Vécsey and Ágnes Cseh for proofreading. The statement of the fundamental theorem of tonality was conjectured by Máté Vécsey.

References

1. Bach, Johann Sebastian. *Vierstimmige Choralgesänge*. Editor: Imre, Sulyok. Editio Musica Budapest, 1982.
2. Benson, Dave. *Music: A Mathematical Offering*. University of Aberdeen, 2008.
3. Kesztlér, Lőrinc. *Összhangzattan* (Harmony). Music theory coursebook for Hungarian high schools. Editio Musica Budapest, 1952.
4. Ligeti, György. Decision and automatism in the piece *Structure 1a* of Pierre Boulez. (Original German title: Entscheidung und Automatik in der *Structure 1a* von Pierre Boulez.) *die Reihe* 4 (1958), p. 38–63. *GS* 1, p. 413–446.
5. Mazzola, Guerino. Mathematical Music Theory, Status Quo 2000. Conference lecture, ETH Zürich, Departement GESS, and Universität Zürich, Institut für Informatik, 15th October 2001. Online reference: <http://recherche.ircam.fr/equipes/repmus/mamux/documents/status.pdf>.
6. Rimsky-Korsakov, Nikolai. *Practical Manual of Harmony*. Carl Fischer, LLC, 2005.
7. Tymoczko, Dmitri. The Geometry of Musical Chords. In: *Science* 313, 72 (2006).
8. Tóbiás, András. *A klasszikus összhangzattan axiomatikája (The axiom system of classical harmony)*, Bachelor thesis in Mathematics, thesis supervisor: Ákos G. Horváth. Budapest University of Technology, 2014. Online reference: math.bme.hu/~tobiasaj/analosszhang_szakdolgozat.pdf
9. Thorpe, Christopher A. *C.P.U. Bach: Using Markov Models for Chorale Harmonization*. Harvard College, Cambridge, Massachusetts, 1998.